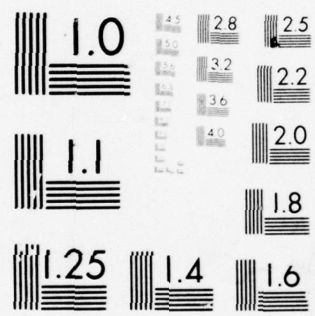
AD-A037 482 CALIFORNIA UNIV RIVERSIDE F/G 12/1
COVARIANCE ANALYSIS OF SOME NONSTATIONARY TIME SERIES.(U)
DEC 76 M M RAO F33615-74-C-4009
UNCLASSIFIED AFFDL-TR-76-128 NL

1 OF 1
AD
A037482

END
DATE
FILMED
4-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-

AFFDL-TR-76-128

12

COVARIANCE ANALYSIS OF SOME NONSTATIONARY TIME SERIES

UNIVERSITY OF CALIFORNIA
RIVERSIDE, CALIFORNIA 92502

DECEMBER 1976

TECHNICAL REPORT AFFDL-TR-76-128
FINAL REPORT FOR PERIOD 1 AUGUST 1974 - 1 AUGUST 1976

Approved for public release; distribution unlimited

AIR FORCE FLIGHT DYNAMICS LABORATORY
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Information Office (IO) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

H. Leon Harter

H. LEON HARTER
Mathematician
Applied Mathematics Group
Analysis & Optimization Branch

FOR THE COMMANDER

HOWARD L. FARMER, Colonel, USAF
Chief, Structural Mechanics Division

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

AIR FORCE - 23 MARCH 77 - 400

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 18 AFFDL-TR-76-128	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) 6 COVARIANCE ANALYSIS OF SOME NONSTATIONARY TIME SERIES.		5. TYPE OF REPORT & PERIOD COVERED Technical - Final technical rept. 1 Aug 1974 - 1 Aug 1976
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 M. M. Rao		8. CONTRACT OR GRANT NUMBER(s) 15 F33615-74-C-4009 Mod. P 00003
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of California Riverside, California 92502		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 12 02 16 7071-02-10
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Flight Dynamics Labs./FBRD Wright-Patterson Air Force Base, Ohio 45433		12. REPORT DATE 11 December 1976
		13. NUMBER OF PAGES 46
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 50p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonstationary time series, covariance analysis, spectral distribution, trend and seasonal models, almost harmonizable time series, correlation characteristic.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Some classes of nonstationary time series, of both continuous and discrete parameter, have been analyzed essentially based on their correlation structure. These classes include the important family, isolated by Kampé de Fériet and Frenkiel and independently by Parzen, and some unstable series. The behavior of the trajectories have been studied in all these cases, which consider the difference schemes as well as the differential equation systems. next page

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

072 350

LB

20. ABSTRACT (Continued)

Both the potential applications and new problems arising from this analysis have been pointed out throughout.

PREFACE

This final report was prepared for the Air Force Flight Dynamics Laboratory, by M. M. Rao under Project 7071, "Research in Applied Mathematics," Task 02, Work Unit 10, "Inference on Multiple Time Series." The work of M. M. Rao was accomplished under under Contract F 33615-74-C-4009 Mod. P 00003. This effort was technically monitored by P. R. Krishnaiah and H. L. Harter of the Air Force Flight Dynamics Laboratories.

TABLE OF CONTENTS

SECTION I

INTRODUCTION AND OUTLINE

In recent times there have been many attempts at extending the analysis of stationary time series to classes of nonstationary cases. The fundamental spectral analysis, so useful for the stationary series, does not play such a central role in the general study. The necessary tools and methods for an analysis of the latter are different and altogether less refined, and still appear more complicated. However, from the point of view of applications, the analysis of nonstationary time series is perhaps more realistic. Consequently, some results based on certain asymptotic considerations related to a 'correlation characteristic' (to be defined below) will be presented in this paper. Also included is some other work on classes of linear stochastic (or autoregressive) equations in discrete as well as continuous time. An outline of the results will be useful here, since it gives the reader a bird's-eye view of the treatment.

Instead of totally abandoning the spectral point of view, Kampé de Fériet and Frenkiel in 1959 have, in a remarkable paper, introduced a class of nonstationary time series (to be called class (KF) hereafter) for its covariance analysis (cf. [1]). There they studied in considerable detail a model which is of the form signal plus noise, where the signal is a time series with zero means and a periodic covariance and the noise is a stationary series with zero mean. A detailed numerical study was made to illustrate the usefulness of this class. Also only slightly later, but independently, Parzen [2] has considered such a model briefly and termed it "asymptotically stationary." This class (KF) has been discussed in some detail by Bhagavan [3] in his recent dissertation. There one of his main results asserts that the class of harmonizable time series belongs to class (KF), and then an ergodic theorem was obtained. In Section II below, the class (KF) has been further analyzed and a more inclusive (new) <u>almost harmonizable class</u> was introduced and shown to be of class (KF). This contains the harmonizable case, but the generalization gives a better insight, and shows the simplicity of class (KF). The consistency of an estimator of the averaged mean function and a weak law of large numbers are established. It is of interest to remark that the classical spectral theory of stationary time series again plays a role here---perhaps justifying the terminology of Parzen's noted above. There are some relations between a class of series introduced by Cramér in (1951), (cf. [4]), called class (C), and (KF) though either does not include the other. A comparison of these approaches is also expounded here, and these relations reappear at many places in the rest of the paper.

Section III is devoted to the covariance analysis of time series governed by difference schemes with not necessarily constant coefficients. Such equations are of interest in treating trend and seasonal variations in various situations which are typically (strongly) nonstationary (cf. Hurwicz [5]). Under some conditions on these coefficients, it is shown that such generated series belong to class (KF) . In general many do not. Some of the latter cases have been analyzed in detail. In order to understand the dependence, approximate recursion equations for sample covariances, which are the sample analogs of the ancient Yule-Walker equations, are presented when the coefficients have linear "time trend." For a descriptive study, correlograms (= the graph of $(k, \rho_t(k))$ where $\rho_t(k)$ is the correlation of X_t with X_{t+k}) and certain approximations are considered. Even here, the computations become involved, but what may be expected in higher order schemes is evidenced. Some other results on estimation and limit behavior of normalized sums are also included.

The continuous time analog of the preceding work refers to the behavior of flows. This is considered in some detail in (the final) Section IV. The corresponding schemes are stochastic differential equations whose coefficients are functions of time. The importance of a class of these schemes in industrial applications has been reported by Hartley [6]. If the coefficients are constants, they represent the motion of a simple harmonic oscillator, driven by random (or white noise) disturbance, and such a model has already been discussed in 1943 (and in earlier classical studies) in the important long article of Chandrasekhar [7]. The latter type of equations have been analyzed by Dym [8], and classified. The general time dependent case is much more involved, and some properties of such an equation are studied. Several results from the theory of ordinary differential equations have special interest here. The basic sample function continuity of the solutions and the conditions under which they belong to class (KF) and the fact that they always belong to class (C), among others, are established. A specialization of the case when the coefficients are constants is illustrated, and the correlogram is analyzed for its asymptotic behavior.

The results presented show how new problems must be attacked for a more complete understanding of these time series. In particular, the classification of the solutions, analogous to [8], will be very interesting. Some aspects of this for vector valued time series have been already given by Goldstein [9], but the sample function behavior when the disturbance is white noise, as in [8], has not been done. The feasibility of such a study is strongly indicated in the present work. Similarly, many other connections and scenic byways are noted but not pursued. Hopefully such work will be considered in the future.

It is recognized that, especially in time series, no result can be taken without adequate demonstration (or at least the explanations that can easily be made precise). For this reason essentially all proofs appear along with the statements of results if an adequate reference is not at hand. Consequently, readers primarily interested in the results are advised to skip the proofs and proceed with the statements, discussions, and remarks of the paper.

SECTION II

A GENERAL CLASS OF NONSTATIONARY TIME SERIES

Let $X = \{X_t, t \in \mathbb{R}\}$ be a second order (real or complex) time series with zero mean, and covariance $K(s,t) = E(X_s \bar{X}_t)$. (The expectation E is on a fixed probability space on which X_t is defined.) Suppose that the $K(\cdot,\cdot)$ satisfies the following condition:

$$\text{(KF)} \qquad r(h) = \lim_{T\to\infty} \frac{1}{T} \int_{|h|/2}^{T-\frac{|h|}{2}} K(s-\tfrac{h}{2}, s+\tfrac{h}{2})\, ds$$

$$= \left(\lim_{T\to\infty} \frac{1}{T} \int_0^{T-|h|} K(s, s+|h|)\, ds = \lim_{T\to\infty} r_T(h) \right), \quad h \in \mathbb{R} ,$$

where the limit is assumed to exist. This condition was introduced by Kampé de Fériet and Frenkiel [1]. The class of time series X satisfying the above condition will be called the class (KF) , and it is analyzed in detail in what follows. It will become clear that this class (KF) is sufficiently general and is very useful in applications. The interest in this definition stems from the positive definiteness of r_T and r , as stressed in [1].

It is clear that, if X is (real and) stationary, so that $K(s,t) = K(t-s)$, then $r(h) = K(h) = K(|h|)$. Thus every stationary process is in (KF) . If $X = Y+Z$, where $Z = \{Z_t, t \in \mathbb{R}\}$ is stationary (with zero mean) and $Y = \{Y_t, t \in \mathbb{R}\}$ has zero mean and a periodic covariance (i.e. $K_Y(s+h_0, t+h_0) = K_Y(s,t)$ for some "period h_0 ") $E(Y_t \bar{Z}_t) = 0$, $t \in \mathbb{R}$, then also

$X \in$ class (KF) , but now X is evidently not stationary. In this representation, the output X is composed of the "signal" Y and "noise" Z which are mutually uncorrelated, and the model describes a communication channel. An elementary example of the Y-process is the following:

$$Y_t = \alpha \cos 2\pi t \ , \quad t \in \mathbb{R} \ , \tag{1}$$

where $E(\alpha) = 0$, $E(\alpha^2) = \sigma^2 > 0$. Then Y is nonstationary and $K_Y(s,t) = \frac{\alpha^2}{2}[\cos \pi(t+s) - \cos \pi(t-s)]$. In this example $r_T(h)$ of (KF) is given by $r_T(h) = K_Z(h) + \frac{\alpha^2}{2} \cos \pi h + o(T)$ so that $X \in$ class (KF) . Also if $K_X(t,t+h) \to a(h)$ as $t \to \infty$ for each h where $-\infty < a(h) < \infty$, then (by the L'Hospital rule) it follows that $X \in$ class (KF) . Another example of X in (KF) is the important nonstationary class called the <u>harmonizable</u> time series. Namely, if the covariance of X is denoted by K_X , then it is representable as:

$$K_X(s,t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{is\lambda - it\mu} d^2\gamma(\lambda,\mu) \tag{2}$$

where $\gamma(\cdot,\cdot)$ is a covariance function of bounded variation on the complex plane (or on the square $(-\pi,\pi] \times (-\pi,\pi]$ if the time series is of discrete parameter). K_X is then called a harmonizable covariance. That this series X is in class (KF) is nontrivial and it is one of the main results of the dissertation [3]. There is a more inclusive class of Cramér, to be called class (C) , generalizing (2); it will be recalled here for comparison with (KF) and for later use.

The time series $X = \{X_t, t \in \mathbb{R}\}$ is said to be of <u>class</u> (C), if it has mean zero and covariance $\tilde{K}_X$, representable as:

$$\tilde{K}_X(s,t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(s,\lambda)\overline{g(t,\mu)}d^2\gamma(\lambda,\mu) \tag{3}$$

where γ is a covariance function of bounded variation on each finite domain of the complex plane, and $\{g(t,\cdot), t \in \mathbb{R}\}$ is a family such that the integral in (3) exists. $\tilde{K}_X$ is then called a covariance of <u>type</u> (C). Thus if γ is of bounded variation, $g(s,\lambda) = e^{is\lambda}$, then (3) reduces to (2). The problem now is to find conditions on g and γ in order that X of class (C) is in class (KF). The class (C) has been analyzed from the point of integral representation by the author [10], and considerable information is available for this family.

For comparison, it will be useful to state the above-noted result from [3] in the following form. It is seen to be included in a more general result proved next.

<u>Proposition</u> 2.1. Let K be a continuous harmonizable covariance. Then

$$R(h) = \lim_{T\to\infty}\frac{1}{T}\int_0^{T-|h|}K(t,t+|h|)dt\ , \qquad h \in \mathbb{R}\ , \tag{4}$$

exists and $R(\cdot)$ is a stationary covariance so that

$$R(h) = \int_{-\infty}^{\infty} e^{ih\lambda}dF(\lambda)\ , \qquad K \in \mathbb{R}\ , \tag{5}$$

for a unique distribution F ($F(+\infty)-F(-\infty) = R(0)$). Moreover,

$$F(\lambda) = \gamma(\infty,\infty) - \gamma(-\lambda,-\lambda) \tag{6}$$

where γ is the covariance given in (2).

In the discrete case there is a corresponding formula (analogous to (4)-(6)), but the relation (6) can be given a more detailed form using the possible discontinuities of γ. (It was treated in [3].) The proof of this result uses estimates of trigonometric sums for integrals on bounded rectangles arising from $\{e^{it\lambda}, t \in \mathbb{R}\}$. In the case of (3), it is clear that one has to restrict the family $\{g(t,\cdot), t \in \mathbb{R}\}$ regarding its growth in relation to γ. In general, the limit (4) for K's of type (C) need not exist. Thus the class (C) is not included in class (KF) nor is the class (KF) contained in the class (C). (See the example in remark following Theorem 4.2.) Some interesting conditions on g will be obtained so that those time series are in class (KF).

Let g of (3) be a bounded (jointly) continuous function. It will now be shown that, if $g(\cdot,\lambda)$ is almost periodic for almost all λ (in particular, $g(t,\lambda) = e^{it\lambda}$ is automatically included), then the corresponding class (C) covariance actually is in (KF) so that (4) and (5) are implied. It is necessary to recall the definition of almost periodicity of g, depending on the parameter λ, to demonstrate the preceding statement.

Definition. Let $D \subset \mathbb{R}^n$ be an open (or a compact) set. A continuous complex function f on $\mathbb{R} \times D$ is said to be almost periodic (a.p.) on $\mathbb{R}$ uniformly relative to D if for each compact set $S \subset D$, $f(\cdot,x)$ is almost periodic for each x in S, i.e., for any $\varepsilon > 0$, and each compact set $S \subset D$, there is a number $\ell_0 = \ell_0(\varepsilon,S) > 0$, such that each interval $I \subset \mathbb{R}$ of length ℓ_0 contains a number $\tau \in I$ for which

$$|f(t+\tau,x)-f(t,x)| \le \epsilon \ , \qquad t\in\mathbb{R}, x\in S \ . \tag{7}$$

The τ is called an ϵ-translation number of f .

It can be shown that the set of all a.p. functions depending on a parameter, satisfying (7), forms an algebra, and if $|f(t,x)| \ge a_S > 0$ for $x \in S \subset D$, and all $t \in \mathbb{R}$, then $\frac{1}{f}$ is also an a.p. function of the same kind. Thus the set $\{e^{it\lambda}, t\in\mathbb{R}, \lambda\in\mathbb{R}\}$ is included in the above, and in fact properly. For an exposition of this class, the reader may consult Yoshizawa ([11], Chapter I). If $D = \{t\}$ is a single point, then this definition reduces to the classical concept of a.p. functions of Bohr. Also observe that an a.p. function is only locally (i.e. on bounded intervals) integrable. In fact, it is bounded for each $x \in S \subset D$.

The following is the desired generalization.

Proposition 2.2. Let K be a covariance function of type (C) , i.e. one which satisfies (3). Suppose that $g(\cdot,\lambda)$ of the integrand in (3) is almost periodic uniformly relative to $\lambda \in D = \mathbb{R}$ and with γ as the covariance of bounded variation. Then the time series $X = \{X_t, t\in\mathbb{R}\}$ with zero mean and covariance K belongs to class (C)$\cap$class (KF) . More precisely,

$$R(h) = \lim_{T\to\infty} \frac{1}{T}\int_0^{T-|h|} K(t,t+|h|)dt = \lim_{T\to\infty} R_T(h) \ , \quad h\in\mathbb{R} \ , \tag{8}$$

exists and defines a stationary covariance on $\mathbb{R}$.

Remark. If $g(t,\lambda) = e^{it\lambda}$, then the hypothesis of this proposition is satisfied by that of the preceding one so that the main result (4) is a consequence of (8). Using the special form of this g, it is possible to obtain (6) connecting γ and the spectral distribution F of R. However, in the present generality, such a relation as (6) is much more involved.

Proof. By symmetry it suffices to consider $h \geq 0$. Now substituting (3) into R_T above and interchanging the integrals (this is obviously legal)

$$R_T(h) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{T}\int_0^{T-h} g(t,\lambda)\overline{g(t+h,\lambda')}dt\, d^2\gamma(\lambda,\lambda') . \tag{9}$$

If $S \subset \mathbb{R}$ is any compact set, and $g(\cdot,\lambda)$, $\lambda \in S$, is a.p., then it follows that $g(\cdot,\lambda)\overline{g(\cdot+h,\lambda')}$, $(\lambda,\lambda') \in S \times S$, is also a.p. So for any fixed but arbitrary h, one has by a classical result (cf. Besicovitch [12], p. 15),

$$\lim_{T\to\infty} \frac{T-h}{T} \lim_{T\to\infty} \frac{1}{T-h}\int_0^{T-h} g(t,\lambda)\overline{g(t+h,\lambda')}dt = a(h;\lambda,\lambda') , \tag{10}$$

exists uniformly in h. But it is clear that $a(h;\lambda,\lambda')$ is bounded for all $h \geq 0$, $(\lambda,\lambda') \in S \times S$ since each a.p. function is bounded. So from (9) and (10), together with Dominated Convergence, one gets

$$R(h) = \lim_{T\to\infty} R_T(h) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a(h;\lambda,\lambda')d^2\gamma(\lambda,\lambda') . \tag{11}$$

But for each T, $R_T(\cdot)$ is clearly positive definite and hence so is $R(\cdot)$. Thus $R(\cdot)$ is a covariance. However, due to the compactness of S and the uniformity involved in (10), it follows that $a(\cdot;\cdot,\cdot)$ on $\mathbb{R} \times S \times S$ is a continuous complex function. From this one easily concludes that $R(\cdot)$ is a (continuous) stationary covariance and then the representation (5) is just the classical

Bochner's theorem. This completes the proof.

Comments. 1. Evidently one may prove an analogous result if the time series is of discrete parameter. It should be observed that the long argument of the harmonizable case of Bhagavan (cf. [3], pp. 72-77) is really a specialized version of the existence of the limit (10) and for the special $g(t,\lambda) = e^{it\lambda}$ (the "characters" of $\mathbb{R}$) the mean value is given by the values of (10) on the diagonal (cf. [12], p. 16, no. 4°) so that the simplifications for (6) result. In the general case (6) no longer holds. Now X is mean continuous, i.e. $E(X_t - X_s)^2 \to 0$, as $s \to t$, so one has only:

$$R(h) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a(h;\lambda,\lambda')d^2\nu(\lambda,\lambda') = \int_{-\infty}^{\infty} e^{ihx}dF(x), \quad h \in \mathbb{R}, \qquad (12)$$

by the Bochner theorem. This F will be called the associated spectral distribution of X .

2. The time series $X = \{X_t, t \in \mathbb{R}\}$ whose covariance K is of type (C) (and $E(X_t) = 0$) for an a.p. function $g(\cdot,\lambda)$, uniformly relative to $\mathbb{R}$, may and will be called an almost harmonizable series. This clearly includes the harmonizable case. Under this generalization, one may profitably consider the more inclusive Besicovitch functions g ($= B^2$- a.p. of [12]), a.p. uniformly relative to $\mathbb{R}$, since for such functions the desired limit (10) again exists. This follows from ([12], p. 93), where one uses the fact that such g's form an algebra and then Lemma 4 is applied there in a slightly modified form (and consequently the arguments on pp. 14-15 of [12] hold). This

extension is necessary to show, for instance, that the Brownian motion (which is not harmonizable) is covered by the above. That the latter is in class (KF) is easy to check directly as observed in [1]. Similarly the Ornstein-Uhlenbeck process is in (KF). Both these are now almost harmonizable (g is not continuous, so B^2 a.p. is needed!). It will be interesting to analyze this set which is contained in class (KF)$\cap$class (C). Here one should perhaps also observe that, if g is B^2-a.p. so that $a(h;\lambda,\lambda')\to 0$ as $|\lambda|+|\lambda'|\to\infty$ sufficiently rapidly, then γ need only be of bounded variation on each compact rectangle of the complex plane.

3. It may also be remarked that g of (3) can be more general than that noted above for the existence of the limit in (10). For instance, if g is locally (i.e. on compact sets) square integrable, then $g(\cdot,\lambda)\overline{g(\cdot+h,\lambda)}$ will be $(c,1)$ summable (i.e., in the sense of first arithmetic mean of Cesàro) while γ is of bounded variation. Many good sufficient conditions are available for it in the literature. This shows that class (KF)$\cap$class (C) contains even the almost harmonizable family as a proper subset.

One of the key applications of the above result is in obtaining conditions for the weak or strong law of large numbers. In a different terminology, this is equivalent to estimating the mean (or the average of the mean function) of the time series X consistently. These problems are natural analogs of the well known stationary theory (cf. Doob [13], pp. 529-530). The following is such an extension and it is substantially due to Bhagavan [3].

Proposition 2.3. Let $X = \{X_t, t \in \mathbb{R}\}$ be an almost harmonizable time series (which is mean continuous). Suppose that its mean function m has the property that $a_0 = \lim_{T\to\infty} \frac{1}{T}\int_0^T m(t)dt$ exists where $m(t) = E(X_t)$. If $\hat{m}_T = \frac{1}{T}\int_0^T X_t dt$ (the sample path integral), then $\lim_{T\to\infty} E(\hat{m}_T - a_0)^2 = F(0+) - F(0-)$, where F is the associated spectral distribution of X (cf., (12)). In particular, if F is continuous at 0, then $\hat{m}_T$ is a strongly consistent estimator of a_0 (or the series X obeys the weak law of large numbers when the limit a_0 exists).

Proof. Let $a_T = \frac{1}{T}\int_0^T m(t)dt$, so that $a_T \to a_0$ as $T \to \infty$. Then writing $K(s,t) = \mathrm{Cov}(X_s, X_t)$, the covariance, and noting $E(\hat{m}_T) = a_T$, one has:

$$
\begin{aligned}
E(\hat{m}_T - a_0)^2 &= E[(\hat{m}_T - a_T) + (a_T - a_0)]^2 = 2[E(\hat{m}_T - a_T)^2 + (a_T - a_0)^2] \\
&= \frac{2}{T^2}\int_0^T\int_0^T K(s,t)dtds + 2(a_T - a_0)^2, \\
&= \frac{2}{T}\int_{-T}^T R_T(h)dh + 2(a_T - a_0)^2 \qquad \text{(cf. (8))}. \qquad (13)
\end{aligned}
$$

The last term of (13) tends to zero, and the first term may be simplified as follows. If R_T did not depend on T, then the limit of the first term of (13) is the desired result, and it is a classical theorem of Bochner (cf. Cramér [14], p. 25). Since by Proposition 2 (since X is mean continuous), R_T is a continuous positive definite function converging to a continuous positive definite R which then has the representation (12), a simple modification of Bochner's proof establishes the present case

Let F_T and F be the bounded nonnegative nondecreasing functions representing R_T and R, as in (12). Then $R_T(h) \to R(h)$ uniformly for h on each bounded closed interval and these are Fourier transforms of F_T, F which must then be uniformly bounded. Moreover $F_T \to F$ at each continuity point of F by ([14], Thm.11). Thus

$$\frac{1}{2T}\int_{-T}^{T} R_T(h)dh = \int_{-\infty}^{\infty} \frac{1}{2T}\int_{-T}^{T} e^{ith} dh dF_T(t)$$

$$= \int_{-\infty}^{\infty} \frac{\sin Tx}{Tx} dF_T(x) \ . \tag{14}$$

Since $\frac{\sin Tx}{Tx} \to \delta_{0x}$ as $T \to \infty$ (δ_{0x} is the delta function) and $F_T \to F$ the result follows easily. In fact, for each $\varepsilon > 0$,

$$\int_{-\infty}^{\infty} \frac{\sin Tx}{Tx} dF_T(x) = \int_{-\infty}^{-\varepsilon} + \int_{\varepsilon}^{\infty} \frac{\sin Tx}{Tx} dF_T(x) + \int_{-\varepsilon}^{\varepsilon} \frac{\sin Tx}{Tx} dF_T(x) \ . \tag{15}$$

The first two integrals on the right side of this equation go to zero as $T \to \infty$ since $\left|\frac{\sin Tx}{Tx}\right| \le \frac{1}{\varepsilon T}$ and $F_T(x) \le \sup_T F_T(\infty) < \infty$. Let ε be chosen so that $(F(\varepsilon)-F(-\varepsilon)) - (F(0+)-F(0-)) < \varepsilon/2$ and that $\pm\varepsilon$ is a continuity point of F. This can be done since the continuity set of F is everywhere dense in $\mathbb{R}$. But $\left|\frac{\sin Tx}{Tx}\right| \le 1$, and if T is large then $F_T(\varepsilon)-F_T(-\varepsilon)$ differs arbitrarily little from $F(\varepsilon)-F(-\varepsilon)$, and it follows that the last term of (15) differs from $F(0+)-F(0-)$ by less than ε. Since $\varepsilon > 0$ is arbitrary, one concludes that

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} R_T(h)dh = F(0+)-F(0-) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} R(h)dh \ . \tag{16}$$

The proposition now follows from (13) and (16). Note that when F has a jump at 0, the estimator $\hat{m}_T$ (of a_0) is not consistent. This completes the proof.

Remark. Since the limit R and hence F are given only indirectly from R_T, it is desirable to have conditions on $\{R_T, T>0\}$ that will ensure the continuity of F at 0 (or at any given point). This will be true if $R_T(h)\to 0$ as $h\to\infty$ uniformly in T or if $\int_0^\infty |R_T(h)|dh$ is bounded as a function of T. The first condition ensures the statement in brackets (by the classical Riemann-Lebesgue lemma) and the latter gives the continuity of F at 0 as the last part of (16) shows.

In view of the importance of R_T and R above, the immediate statistical problem is to estimate these functions, with the consistency property, at least. (Assume $m = 0$; otherwise one may consider the product moments directly.) Thus the natural estimates are

$$\hat{R}_T(h) = \frac{1}{T}\int_0^{T-h} X_t \bar{X}_{t+h}\, dt = \frac{1}{T}\int_{-\infty}^{\infty} X_t^T \overline{X_{t+h}^T}\, dt, \tag{17}$$

where $X_t^T = X_t$ for $0\le t\le T$, $= 0$ for $t>T$. Under the mean continuity of the X, such that $\int_0^T K^2(t,t)dt < \infty$, the estimator $\hat{R}_T$ of (17) is well defined, $E(\hat{R}_T(h)) = R_T(h)$ ---an unbiased estimator of R_T. To fulfill the consistency condition, i.e. for $X\in$ class (KF), $\hat{R}_T(h)\to R(h)$, $h\in\mathbb{R}$, in mean, one has to assume somewhat more on X. The following sufficient conditions were given by Parzen [2] (cf. also [1]). Thus consider:

(i) for each t, assume $E(|X_t|^4) < \infty$, (ii) if $\rho(t_1,t_2,t_3,t_4) = E(X_{t_1}X_{t_2}X_{t_3}X_{t_4})$ then ρ is Lebesgue integrable on each bounded interval of $\mathbb{R}^4$, and (iii) if $C_t(v) = \int_0^t \mathrm{Cov}(X_s X_{s+v}, X_t X_{t+v})\, dt$ (which exists by (ii)), then $\lim_{t\to\infty} C_t(v) = 0$ for each $v \geq 0$.

If (i)-(iii) hold then $\mathrm{Var}(\hat{R}_T(h)) \to 0$ as $T \to 0$, for each $h \geq 0$, so that $E(\hat{R}_T(h) - R(h))^2 \to 0$. This may be checked by a direct computation (cf. also [2]). In [1] an interesting example of a periodic covariance of a time series consisting of symmetric bounded random variables is given which satisfies the above conditions. The accuracy of the approximation of the estimator with R was then illustrated by a numerical example. The reader is referred to this instructive case in [1] to gain an insight into the generality of the class (KF) of nonstationary time series.

The next two sections will now be devoted to another class of nonstationary series (= processes) generated by certain stochastic difference and differential equations for which a different set of methods will be needed. The latter are related to the "correlogram analysis," and these will be discussed.

SECTION III

NONSTATIONARY SERIES GENERATED BY DIFFERENCE EQUATIONS

(a) Motivation. If $X = \{X_t, t\in\mathbb{R}\}$ is a stationary time series, then the correlation function $\rho(\cdot)$ is given by $\rho_t(h) = \rho(h) = \mathrm{Cov}(X_t, X_{t+h})/\mathrm{Var}\, X_t$, and is independent of t. As a function of h, i.e. of the lag, $\rho(\cdot)$ is taken as an indicator of the dependence of X_t on X_{t+h} for large h. Clearly $\rho(h) = \overline{\rho(-h)}$, $|\rho(h)|\le 1$. Suppose that X is nonstationary, but is in class (KF) In this case $\rho_t(\cdot)$ depends on t, but for any $a>0$, and $h\in\mathbb{R}$,

$$\lim_{T\to\infty} \tilde{\rho}_{T,a}(h) = \lim_{T\to\infty} \frac{\frac{1}{T-a}\int_a^{T-h} K(t,t+h)dt}{\frac{1}{T-a}\int_a^{T-h} K(t,t)dt} = \frac{R(h)}{R(0)} = \bar{\rho}(h) \quad \text{(say)}. \tag{1}$$

Here R is the same as in (8) of the preceding section with $K(s,t) = \mathrm{Cov}(X_s, X_t)$. Since $R(\cdot)$ is a stationary covariance by virtue of the fact that $X\in$ class (KF), $\bar{\rho}(\cdot)$ does not depend on t or a. Further, several properties of the time series X are reflected in the behavior of R and hence of $\bar{\rho}$. In case that X is stationary then $\bar{\rho} = \rho$. On the other hand, if $\lim_{t\to\infty} K(t,t+h) = a(h)$ exists (which is stronger than being in (KF)), then the limit of (1) exists and one gets the same $\bar{\rho}(\cdot)$. This is a consequence of known work in classical analysis. Motivated by this observation, one may consider the behavior of $\tilde{\rho}$ for a class of nonstationary time series where

$$\tilde{\rho}_t(h) = \frac{\mathrm{Cov}(X_t, X_{t+h})}{\mathrm{Var}\, X_t}, \tag{2}$$

called hereafter the correlation characteristic of X. In general

$\tilde{\rho}$ is not a correlation function, but asymptotically its behavior is that of the latter (because of (1)) at least for the series of class (KF). In this section the time series governed by certain difference (and in the next section differential) equations will be studied. Some of the following material (in both sections) appears in a very tentative form in the author's early study [15].

(b) A setting of the problem. Suppose that the output of a noisy communication channel follows a linear model of signal plus noise type as follows. The output X_t at time t depends on the immediate past up to k units linearly, and then a white noise disturbance enters. Thus

$$X_t = \sum_{i=1}^{k} a_i(t) X_{t-i} + \varepsilon_t = S_t + \varepsilon_t\,, \qquad t \geq 1\,, \tag{3}$$

where $a_i(\cdot)$ are some (non-stochastic) functions of time specified by the type of channel. *This model may also* be used for a dynamic economic situation where X_{t-i}'s are called "lagged values" of the "endogenous" variable X_t or sometimes "exogenous" variables also. Thus the signal S_t is a linear function of the past k terms and $\{\varepsilon_t, t \geq 1\}$ are assumed to be independent identically distributed variables with zero mean and a finite variance ($= \sigma^2$, say). Now to study the properties of X_t, the difference equation (3) may be solved by classical methods (due to D. André, 1878) and the solution is given by the expression:

$$X_t = \sum_{i=0}^{t-1} \varphi(t,i) \varepsilon_{t-i} + \sum_{i=0}^{t+k-1} \varphi(t,i) c_{t-i}\,, \qquad t \geq 1\,, \tag{4}$$

where $X_i = c_i$, $i = -k+1, \ldots, 0$, are the constant initial values, and

$$\varphi(t,m) = \sum_{k_1+\cdots+k_i=m} \prod_{j=1}^{i} a_{k_j}(t-\sum_{r=0}^{j-1} k_r) \ , \ 0\le m\le t, k_0=0 \ , \tag{5}$$

the sum ranging over all partitions of m into integers k_i (≥ 0) (cf., Jordan [16], p. 588). For instance, if $c_i = 0$ and $a_i(t) = \alpha_i$, a constant, then the complicated looking expressions (4) and (5) reduce to familiar forms. To see this, let the characteristic equation of the difference equation (3), namely,

$$\lambda^k - \alpha_1\lambda^{k-1} - \cdots - \alpha_k = 0 \tag{6}$$

have simple roots $\lambda_1,\ldots,\lambda_k$. Then (5) becomes

$$\varphi(t,m) = \varphi(t-m) = \sum_{j=1}^{k} \beta_j\lambda_j^{t-m} \ , \qquad 1\le m\le t, \sum_{j=1}^{k} \beta_j = 1 \ . \tag{7}$$

The β_j's further satisfy

$$\sum_{j=i}^{k} \beta_j\lambda_j^{t-1} = 0 \ , \qquad t=0,-1,\ldots,-k+2 \ . \tag{7'}$$

For a discussion of this case, see Mann and Wald [17] and the author ([18], p. 330). Let us now specify conditions on the model in order that the time series $X = \{X_t, t\ge 1\}$ be in class (KF) , and point out instances when it is not in class (KF) but for which $\tilde{\rho}_t$ of (2) has a limit as $t\to\infty$.

Since for an asymptotic study the initial values are not important, set $c_i = 0$, $i=0,-1,\ldots,-k+1$ in (4). Suppose further that the $a_i(\cdot)$ satisfy

$$\sum_{m=0}^{t} |\varphi(t,m)|^2 \le M < \infty \ , \qquad t\ge 1 \ . \tag{8}$$

Then from (4), since $E(X_t) = 0$, one finds

$$r(s,t) = E(X_s\bar{X}_t) = \sigma^2 \sum_{m=0}^{s} \varphi(s,m)\overline{\varphi(t,m)}\ , \qquad 1\le s\le t\ ,$$

$$= \sigma^2 \sum_{m=0}^{s} \varphi(s,m)\overline{\varphi(s+h,m)}\ . \qquad h=t,s\ .$$

It follows from (8) that $|r(s,s+h)|\le M$ for all h, and $s\ge 1$ (by the Cauchy inequality). Hence, for each $h\ge 0$,

$$\lim_{N\to\infty} \frac{N-h}{N}\cdot\frac{1}{N-h}\sum_{s=1}^{N-1} r(s,s+h) = R(h) \tag{9}$$

exists (by the (c,1)-summability method), and by symmetry for all h. This implies the following statement.

<u>Proposition</u> 3.1. If a second order time series $X = \{X_t, t\ge 1\}$ is generated by the equation (3), and the $a_i(\cdot)$'s satisfy (8), then the nonstationary series X belongs to class (KF).

It should be observed that the condition (8) does not involve any mention of the roots of (6) even when the a_i's are constants. To understand the significance of (8), it is useful to specialize. Let the a_i be constants and suppose the roots of (6) lie inside the unit circle of the complex plane and are distinct. Let $\delta = \max_j |\lambda_j|$ so that $\delta < 1$. Then (8) is automatic since

$$\sum_{m=0}^{t} |\varphi(t,m)|^2 = \sum_{m=0}^{t} |\varphi(t-m)|^2 = \sum_{m=0}^{t}\left|\left[\sum_{j=1}^{k} \beta_j\lambda_j^{t-m}\right]\right|^2 \text{, by (7),}$$

$$\le \sum_{j,j'=1}^{k} |\beta_j\beta_{j'}|\left(\frac{1-(\lambda_j\bar{\lambda}_{j'})^{t+1}}{1-\lambda_j\bar{\lambda}_{j'}}\right)$$

$$\le \sum_{j,j'=1}^{k} |\beta_j\beta_{j'}|\frac{1+\delta^{2(t+1)}}{1-\delta^2} \le \frac{2}{1-\delta^2}\sum_{j,j'=1}^{k} |\beta_j\beta_{j'}| = M_0 < \infty\ .$$

Thus the time series generated by (3) with constant coefficients having all the roots of its characteristic equation distinct and lying inside the unit circle, belongs to class (KF). On the other hand, if at least one of the roots of (6) is on or outside the unit circle (i.e. $\delta \geq 1$) then the resulting time series generated by (3) is nonstable or explosive and will not be in class (KF). This will now be illustrated by a family of time series, investigated in the literature.

Thus in the constant coefficient case of (3), let $\delta > 1$, $\delta = \max_j |\lambda_j|$ of (6). Then under the same (remaining) hypothesis as above,

$$E(|X_t|^2) = \sum_{t=0}^{t} |\varphi(t,m)|^2 = O(\delta^{2t}) , \tag{10}$$

so that (8) is violated strongly. This was shown by the author ([19], Lemmas 8, 15). Moreover, by ([19], Lemma 9) one finds

$$\tilde{\rho}_t(h) = \frac{\mathrm{Cov}(X_t, X_{t+h})}{\mathrm{Var}\, X_t} \to \delta^{-(h-2)} \qquad (\neq 0), t \to \infty . \tag{11}$$

A similar conclusion holds if $\delta = 1, k = 1$, for (10) ($E(|X_t|^2) = O(t^2)$). These two cases imply that the limit demanded for (KF) of Section II cannot exist, and so these time series do not belong to class (KF). This example may be taken as a further justification of Parzen's term "asymptotically stationary" for the series in classes like (KF) (cf. [2]).

The variable coefficient case of (3) was found to be of interest in some meteorological applications (cf. [5]). Consequently, a class of simple variable coefficient schemes will be investigated in some detail below. They will throw some light on the behavior of the resulting time series.

(c) A first order model. The correlation characteristic, given by (2), is designed to reflect the dependence of X_t on X_{t+h} for large t and h. It was noted that this behaves asymptotically (as $t\to\infty$) as a correlation function for members of class (KF). How does it behave in the "explosive" cases? Equation (11) gives an indication of the constant coefficient case. Here the variable coefficient model of first order will be considered, for a striking illustration. This is also the one proposed in [5], describing a times series with linear trend, namely,

$$X_t = a_t X_{t-1} + \varepsilon_t , \quad a_t = a_0 + t a_1 , \qquad t\geq 1, a_0\neq 0 . \tag{12}$$

The series $\{\varepsilon_t, t\geq 1\}$ consists of uncorrelated random variables with mean zero and variance σ^2, $X_t = 0$, $t\leq 0$ (i.e. the initial values are zero).

The solution of (12) may be obtained by iteration. Alternately, let $p_t = \prod_{i=0}^{t-1} a_i$ so $p_1 = a_0$. If we set $Y_t = X_t/p_t$, (12) becomes

$$Y_{t+1} - Y_t = \varepsilon_t / p_{t+1} , \qquad t\geq 0 . \tag{13}$$

If we write $X_t = Y_t p_t$, the solution of (13) (hence of (12)) is:

$$X_t = \sum_{i=0}^{t} p_t (\varepsilon_i / p_{i+1}) . \tag{14}$$

Hence for $h\geq 0$, the covariance is given by

$$\mathrm{Cov}(X_t, X_{t+h}) = p_t p_{t+h} \sum_{i=0}^{t} \frac{\sigma^2}{p_{i+1}^2} , \qquad t\geq 1 . \tag{15}$$

Since $\operatorname{Var} X_t$ is obtained from (15) for $h = 0$, (2) becomes:

$$\tilde{\rho}_t(h) = p_{t+h}/p_t = \prod_{i=t}^{t+h-1} a_i \ . \tag{16}$$

This can be made arbitrarily large for large h, for appropriate a_0, a_1 (e.g. if $|a_0 + a_1 t| > 1$). On the other hand, the actual correlation $\rho_t(h)$ is:

$$\rho_t(h) = \left(\sum_{i=0}^{t} p_{i+1}^{-2} \Big/ \sum_{i=0}^{t+h} p_{i+1}^{-2} \right)^{\frac{1}{2}} , \tag{17}$$

and $|\rho_t(h)| \to \alpha(h)$ $(\neq 0)$, as $t \to \infty$. Thus for such time series the correlation does not tend to zero for large t and large h. The correlation characteristic $\tilde{\rho}_t(\cdot)$ magnifies this phenomenon. The above computation also implies that the <u>correlogram</u> (= the graph of $(h, \rho_t(h))$, $h \geq 0$, for any t) does not dampen as the lag h increases if the coefficients contain a linear time trend. Since the same character is maintained in using $\tilde{\rho}_t(h)$ instead of $\rho_t(h)$ and since $\tilde{\rho}_t(h)$ is computationally simpler than $\rho_t(h)$, it will be considered in what follows for a structural study of the "explosive" time series. The resulting graph may be called an <u>approximate correlogram</u>.

If the linear trend is replaced by the reciprocal trend, i.e., $a_t = a_0 + \frac{a_1}{t}$, $t \geq 1$, then after a similar but more tedious computation one finds that $\tilde{\rho}_t$ still exhibits nearly the same properties unless more stringent conditions such as (8) are imposed. The details are omitted.

(d) <u>Limiting behavior of normalized sums from class</u> (KF). The behavior of series given by (3), subject to a condition implying

(8), is reasonable in that such series obey the central limit theorem for dependent variables. Let $S_n = X_1+\cdots+X_n$, $X_i=0, i\leq 0$. Suppose that (i) the ε_t's are independent and identically distributed with means zero and variance σ^2, (ii) the $\varphi(t,m) = \varphi(m)$ is independent of t (which is implied by the case that $a_i(\cdot)$ of (3) are constants), and (iii) $\sum_{m=1}^{\infty} |\varphi(m)| < \infty$. Under these conditions the following assertion obtains:

Proposition 3.2. Let the time series $\{X_t, t\geq 1\}$ be generated by (3) and let, moreover, conditions (i)-(iii) hold. Then

$$\lim_{n\to\infty} P[S_n < x\sqrt{\operatorname{Var} S_n}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du . \tag{18}$$

Thus, S_n obeys the central limit law.

This result follows from some known results of Diananda and Anderson (cf. [20], Thm. 7.7.8). In fact, $X_n = \sum_{m=0}^{n} \varphi(m)\varepsilon_{n-m} = \sum_{m=0}^{\infty} \tilde{\varphi}(m)\varepsilon_{n-m}$ where $\tilde{\varphi}(m) = \varphi(m)$, $1\leq m\leq n$, $= 0$ otherwise. Since the ε_m are independent and identically distributed, $\frac{\operatorname{Var} S_n}{n} \to 1$, as a simple computation shows. This is sufficient to invoke the above cited theorem.

The restrictive condition (ii) may be relaxed in some cases. If $a_i(t) = a_i + \tilde{a}_i t^{-2}$, $t\geq 1$, then

$$\varphi(t,m) = \sum_{k_1+\cdots+k_i=m} a_{k_1}(t)a_{k_2}(t-k_1)\cdots a_{k_i}(t-\sum_{j=1}^{i-1} k_j) , \quad k_0=0 ,$$

$$= \sum a_{k_1}a_{k_2}\cdots a_{k_i} - O(\frac{1}{t^2}) = \tilde{\varphi}(m) - O(\frac{1}{t^2}) . \tag{19}$$

Hence

$$X_t = \sum_{m=1}^{t} \tilde{\varphi}(m)\varepsilon_{t-m} + O\left(\frac{1}{t^2} \sum_{m=1}^{t} \varepsilon_{t-m}\right) . \tag{20}$$

One checks that the variance of the last term in (20) is $o(\frac{1}{t^3})$ It has mean zero. So if $S_n = \tilde{S}_n + \tilde{S}'_n$, corresponding to the decomposition of (20), then $\tilde{S}'_n \to 0$ in probability and $\{S_n, n\geq 1\}$ satisfies (iii). Hence, from a form of the classical Slutzky's theorem, S_n and $\tilde{S}_n$ have the same limit distribution. Thus the following result holds:

<u>Proposition</u> 3.3. Let $\{X_t, t\geq 1\}$ be given by the scheme (3). Suppose that conditions (8) and (i) hold, where $a_i(t) = a_i + \tilde{a}_i t^{-2}$ If $\tilde{\varphi}(m)$ of (19) satisfies (iii), then the series obeys the central limit theorem, i.e. (18) holds for this series.

It is clear that other conditions can be formulated on the coefficients $a_i(t)$, to obtain corresponding results. Assuming that the ε_t's have four moments, this result with a sketch of proof was indicated in [15], using Lyapunov's theorem somewhat on the lines of Marsaglia ([21], Thm. 3). However Parzen ([22], p. 254) has a better result on these problems.

(e) <u>Approximate recursion equations for sample covariances</u>. Since the difference equations with coefficients depending linearly on time have been noted to be of interest in describing trend variation (cf. [5]), it will be useful to have recursion formulas for computing sample covariances from data, even though the series is not in class (KF) . Here an account of this problem will be presented for a first order scheme. Similar results for the higher order schemes are much more involved computationally. The recursion equations are just the sample analogs of the ancient <u>Yule-Walker equations</u> of the autoregressive systems (cf. [20], p. 174). These (sample) equations are useful for a descriptive study of time series.

Consider a time series defined by the equation

$$X_{t+1} = (a_0 + a_1 t)X_t + \varepsilon_{t+1}, \qquad t \geq 0, \tag{21}$$

where ε_t's are independent, identically distributed mean zero random variables with finite variance, $X_t = 0$ for $t \leq 0$ so that there are no records of the series for the past and it starts from scratch. Here a_0 and a_1 are unknown parameters and can be estimated by the least squares method. Thus a simple minimization based on N observations gives $\hat{a}_{0N}, \hat{a}_{1N}$ as the estimators of a_0, a_1 by the following equations:

$$\begin{aligned}
\hat{a}_{0N} &= \frac{1}{D_N}\left[\left(\sum_{t=1}^{N} tX_{t+1}X_t\right)\left(\sum_{t=1}^{N} tX_t^2\right) - \left(\sum_{t=1}^{N} t^2X_t^2\right)\left(\sum_{t=1}^{N} X_tX_{t+1}\right)\right], \\
\hat{a}_{1N} &= \frac{1}{D_N}\left[\left(\sum_{t=1}^{N} tX_t^2\right)\left(\sum_{t=1}^{N} X_tX_{t+1}\right) - \left(\sum_{t=1}^{N} X_t^2\right)\left(\sum_{t=1}^{N} tX_tX_{t+1}\right)\right], \\
D_N &= \left(\sum_{t=1}^{N} tX_t^2\right)^2 - (\Sigma X_t^2)\left(\sum_{t=1}^{N} t^2X_t^2\right).
\end{aligned} \tag{22}$$

It is desirable to establish the consistency of these estimators, i.e., to show that $\hat{a}_{iN} \to a_i$, $i=1,2$, in probability as $N \to \infty$. This is true if $a_1 = 0$. The work in ([18], Thm. 5) and [19]) indicates that the general statement is true under some conditions on the distribution of ε_t's. The actual details are nontrivial and will not be considered here. Since the general behavior of the series can be understood to some extent from the properties of the correlation characteristic (or the approximate correlogram of the series) the sample product moment equations will now be derived for this model. This already shows the difficulties involved.

It will be convenient to adopt the following notation, from [5], for simplifications. Thus set (with $X_t = 0, t \le 0$) for $t \ge 0$,

$$Z_{1t} = X_t \ , \ Z_{2t} = tX_t \ , \quad \delta_0 = 1, \delta_1 = -a_0, \delta_2 = -a_1 \ . \tag{23}$$

Then the equation (21) becomes

$$\delta_0 Z_{1t+1} + \delta_1 Z_{1t} + \delta_2 Z_{2t} = \varepsilon_{t+1} \ , \qquad t \ge 0 \ . \tag{24}$$

For the product moments, the following additional abbreviations are seen to be useful.

$$E_k^N = \sum_{t=1}^{N-k} \varepsilon_t \varepsilon_{t+k} \ , \quad S_{ik}^N = \sum_{t=1}^{N-k} Z_{it} Z_{it+k} \ , \qquad i=1,2 \ ,$$

$$S_{12\,k}^N = \sum_{t=1}^{N-k} Z_{1t} Z_{2t+k} \ , \quad S_{21\,k}^N = \sum_{t=1}^{N-k} Z_{2t} Z_{1t+k} \ . \tag{25}$$

Since $\delta_0 = 1$, using (24), one gets after a small computation:

$$\begin{aligned} &\delta_1 S_{1k+1}^N + (1+\delta_1^2) S_{1k}^N + \delta_1 S_{1k-1}^N + \delta_2^2 S_{2k}^N + \\ &\qquad \delta_2 (S_{12k-1}^N + S_{21k+1}^N) + \delta_1 \delta_2 (S_{12k}^N + S_{21k}^N) = E_k^N \ . \end{aligned} \tag{26}$$

But $S_{12k}^N = S_{21k}^N + kS_{1k}^N$. So (26) can be simplified as:

$$\begin{aligned} &\delta_1 S_{1k+1}^N + S_{1k}^N (1+\delta_1 \delta_2 k + \delta_1^2) + S_{1k-1}^N (\delta_1 + (k-1)\delta_2) + \\ &\qquad \delta_2 [S_{21k+1}^N + 2\delta_1 S_{21k}^N + S_{21k-1}^N + \delta_2 S_{2k}^N] = E_k^N \ . \end{aligned} \tag{27}$$

This is the sample analog (for $k \ge 1$) of the Yule-Walker equation of [20] above. Note that $\{E_t^N, N \ge k\}$ is a series of uncorrelated random variables.

If the "initial point" δ_1 of the linear trend is known a priori , then (27) can be recast in a better form. For then, if we set $Y_t = X_t + \delta_1 X_{t-1}$, $\tilde{Y}_t = X_{t-1} + \delta_1 X_t$, $Q_{1k}^N = \sum_{t=1}^{N-k} [\varepsilon_{t+k} Y_{t+1} - \varepsilon_{t-1} \tilde{Y}_{t+k}]$, (27) can be expressed after a small computation as:

$$\delta_1 S_{1k+1}^N + S_{1k}^N (1+\delta_1 \delta_2 k) - S_{1k+1}^N (\delta_1 - (k-1)\delta_2) - S_{1k-2}^N = Q_{1k+1}^N . \tag{28}$$

Treating Q_{1k}^N's as the correlated disturbance, one notes that the product moment equation of the first order scheme (21) (or (24)) is a third order difference equation whose coefficients depend (linearly) on the lag. If $C_k^N = \frac{1}{N-k} S_{1k}^N$, $Q_k^N = \frac{1}{N-k} Q_{1k}^N$, then (28) reduces to an equation of sample covariances:

$$\delta_1 C_{k+1}^N + (1+\delta_1 \delta_2 k) C_k^N - (\delta_1 - (k-1)\delta_2) C_{k-1}^N - C_{k-2}^N = Q_k^N . \tag{29}$$

Taking expectations of this equation, one gets the Yule-Walker relations. If $\delta_2 = 0$ (so there is no trend), then (27) and (29) reduce to a corresponding known (standard) case (cf. [20], p. 124).

Similar considerations with a reciprocal trend (i.e. $a_t = a_0 + \frac{a_1}{t}$ for (21)) lead to the product moment equations, corresponding to (27) or (29), with coefficients depending on the reciprocals of the lag. (This and a second order scheme were considered in [15], but the computations are too long for a treatment here.) From (29) one can easily obtain the correlation characteristic $\tilde{\rho}_N(k) = \frac{C_k^N}{C_0^N}$, and then graph the approximate correlogram. An explicit expression for this can be obtained from (29), by the method indicated in (4)-(5). A formula for C_k^N is given below from (29) without the intermediate computation. Let $f_0 = \delta_1 = -a_0$,

$f_1(k) = 1 + \delta_1 \delta_2 k = 1 + k a_0 a_1$, $f_2(k) = -\delta_1 + \delta_2(k-1) = a_0 - (k-1)a_1$, and $f_3 = 1$. Then

$$C_k^N = \sum_{m=1}^{k} f_0^{-m} (\Sigma f_{t_1}(k_1) \cdots f_{t_i}(k_i)) Q_{k-m} , \tag{30}$$

where (i) $t_1 + \cdots + t_i = m$, (ii) $k_1 = k$, $k_2 = k - t_1$, $k_3 = k - t_1 - t_2$ etc. The graph $\{(k_1, \tilde{\rho}_N(k)), 1 \le k < N\}$, for large enough N , gives an indication of the dependence behavior of the time series described by (21), by the earlier treatment.

SECTION IV

SERIES GENERATED BY DIFFERENTIAL EQUATIONS: FLOWS

(a) Introduction. The preceding analysis leads to the continuous time analog (or the stochastic differential equations case) of the problem involving again nonstationary time series. This is useful both for a comparison with the discrete case above, as well as for an independent study. It also brings up some interesting new problems.

Let $\{X_t, t \in T, T \subseteq \mathbb{R}\}$ be a continuous parameter time series (X_t and $X(t)$ are synonymous) governed by the differential equation:

$$\frac{d^2X(t)}{dt^2} + a(t)\frac{dX(t)}{dt} + b(t)X(t) = \varepsilon(t) \qquad (1)$$

where $\{\varepsilon(t), t \in T\}$ is the white noise disturbance and a, b are real functions on T. By definition of white noise, $\varepsilon(t)$ is the (generalized) derivative of Brownian motion $\{B(t), t \in T\}$; and thus (1) is a symbolic equation which cannot be interpreted in the classical sense of differential calculus. (Such a problem does not, of course, arise in the discrete case.) However, the classical computations carried out formally can be justified (in the integrated form) with the concept of a stochastic integral replacing $\varepsilon(t)dt$ by $dB(t)$, and this will be made precise below. Potential applications of this model abound. Taking $a = 0$, $b(t) = b' + b''t$, Hartley [6] indicated an industrial application and carried out a correlogram analysis of the $X(t)$-series using classical methods (and formal computations). In fact, assuming that the $\varepsilon(t)$ is integrable (in the calculus sense), he has carried out the analysis using "Airy Integrals" and then studied the covariance characteristic. [In a conversation at the IMS meetings in Ames, Iowa in 1957, he mentioned that the Weber differential equation method would be better suited for such a problem.] If a,b are constants, then a related problem was considered by Nagabhushanam ([23], p. 482) where $X(t)$ is called a "primary process" obtained by an inversion. Since the Brownian motion is nondifferentiable in the classical sense so that the $\varepsilon(t)$ of (1) does not exist, a slightly different route will be followed here to validate such an analysis for a solution of (1).

(b) A general second order problem. The method of attack here is quite simple. First consider the problem with a formal manipulation and express the solution in terms of an integral. If we replace $\varepsilon(t)dt$ by $dB(t)$, where $\{B(t), t \in T\}$ is the Brownian motion, and then interpret the integral as the stochastic integral of a nonstochastic (or "sure") function relative to Brownian motion, the solution is rigorously definable. [Of course, the $\varepsilon(t)$-process will not be as general as in [6], but the present assumption will be in force throughout this section. For a rigorous treatment, some such restriction is necessary.]

The differential equation (1), written symbolically as

$$d\dot{X}(t) + a(t)\dot{X}(t)dt + b(t)X(t)dt = dB(t) , \qquad (2)$$

is then regarded as the equation leading to a well-defined solution, where $\dot{X}(t) = \frac{dX}{dt}$. To make this more precise, let $T = [a_0, b_0)$, a bounded interval, and $a(\cdot), b(\cdot)$ be continuous real functions on $\overline{T}$. Set $Q = \dot{X}$, $A(t) = \begin{bmatrix} a(t) & b(t) \\ -1 & 0 \end{bmatrix}$, $W(t) = \begin{bmatrix} B(t) \\ 0 \end{bmatrix}$, $\eta(t) = \begin{bmatrix} \varepsilon(t) \\ 0 \end{bmatrix}$ and $Z(t) = \begin{bmatrix} Q(t) \\ X(t) \end{bmatrix}$. Then (1) may be expressed compactly as:

$$dZ(t) + A(t)Z(t)dt = \eta(t)dt = dW(t) \ . \tag{3}$$

To solve this (vector) differential equation with a standard method in the classical theory of ordinary differential equations, consider the 2-by-2 matrix differential equation associated with the homogeneous part of equation (3), i.e., the (nonstochastic) equation:

$$dY(t) = Y(t)A(t) \ , \qquad t \in T, \ \det(Y(a_0)) \neq 0 \ , \tag{4}$$

where "det" stands for *determinant*. *Then premultiplying* (3) by $Y(t)$ and using (4) one gets

$$\frac{d}{dt}(Y(t)Z(t)) = Y(t)\eta(t)$$

so that formally one has the solution of (3) as:

$$Z(t) = Y(t)^{-1} \int_{a_0}^{t} Y(u)\eta(u)du + Y^{-1}(t)Y(a_0)Z(a_0) \ . \tag{5}$$

The fact that $Y(t)$ satisfying the equation (4), if nonsingular for one $t \in T$, has the same property for all $t \in T$ is used here. This is because (4) implies $\det(Y(t)) = \det(Y(a_0)) \cdot \exp(\int_{a_0}^{t} \mathrm{trace}(A(u))du)$ (cf. Coddington and Levinson [24], p. 28). Thus (5) is well defined and $Z(t)$ is obtained as soon as $Y(t)$ is solved from (4). Since $Y(t)^{-1}$ is bounded for each t, it can be taken inside the integral also. Thus (5) can be expressed as:

$$Z(t) = \int_{a_0}^{t} Y(t)^{-1} Y(u)dW(u) + Y^{-1}(t)Y(a_0)Z(a_0) \ , \qquad (6)$$

where the first term on the right is now rigorously defined as the stochastic integral of the first kind (see Doob [25], p. 352). The uniqueness of the solution is immediate for the given initial condition $Z(a_0)$, because if $\tilde{Z}$ is another solution, then $Z-\tilde{Z}$ will be a solution of the homogeneous equation $\frac{dU}{dt}+A(t)U(t) = 0$, with $U(a_0) = 0$ as the initial condition. This is a nonstochastic equation and the standard theory implies that $U \equiv 0$ is its only solution. Thus $Z = \tilde{Z}$. Hence it remains to find $Y(t)$.

Since (4) is a nonstochastic equation, one can apply the classical Picard method of approximation. Writing $\tilde{A} = A^*$, $\tilde{Y} = Y^*$ ($*$ for transpose) and integrating (4), one finds

$$\tilde{Y}(t) = \tilde{Y}(a_0) + \int_{a_0}^{t} \tilde{A}(u)\tilde{Y}(u)du \ . \qquad (7)$$

Now substituting for $\tilde{Y}$ and iterating, one obtains

$$\tilde{Y}(t) = \tilde{Y}(a_0) + \int_{a_0}^{t} \tilde{A}(t_1)\tilde{Y}(a_0)dt_1 + \cdots + \int_{a_0}^{t} \tilde{A}(t_1) \int_{a_0}^{t_1} \tilde{A}(t_2) \cdots \int_{a_0}^{t_{n-1}} \tilde{A}(t_n)\tilde{Y}(a_0)dt_n \cdots dt_1 + R_n \ , \qquad (8)$$

where

$$R_n = \int_{a_0}^{t} \tilde{A}(t_1) \int_{a_0}^{t_1} \tilde{A}(t_2) \cdots \int_{a_0}^{t_n} \tilde{A}(u)\tilde{Y}(u)du dt_n \cdots dt_1 \ .$$

By hypothesis $a(\cdot)$ and $b(\cdot)$ are continuous on the compact interval $\bar{T}$. So $\|\tilde{A}(t)\| = \sqrt{\mathrm{trace}(\tilde{A}(t)A(t))} \le M < \infty$, $t \in \bar{T}$, and similarly $\|Y(t)\| \le N < \infty$. [Here N is obtained as follows. Let y_1, y_2 be the linearly independent pair of vector solutions of the homogeneous equation of (3). Then these are continuous on $\bar{T}$. If N_1 is the upper bound on the norms of the vectors y_i, $i=1,2$, then $Y = (y_1, y_2)$ in (4) so that $\|Y\| \le \sqrt{2}\, N_1 = N$.] Consequently

$$\|R_n\| \le M^n N \frac{(t-a_0)^n}{n!} \le M^n N \frac{(b-a_0)^n}{n!}, \qquad t \in \bar{T}. \tag{9}$$

Hence $R_n \to 0$ uniformly in t as $n \to \infty$, and the series in (8) converges absolutely and uniformly and defines a solution of (4). In particular, if $\alpha(t,t_0) = \int_{t_0}^{t} \tilde{A}(u)du$, then

$$\int_{a_0}^{t} \tilde{A}(t_1) \int_{a_0}^{t_1} \tilde{A}(t_2) dt_2 dt_1 = \int_{a_0 \le t_2 \le t_1 \le t} \tilde{A}(t_1)\tilde{A}(t_2) dt_2 dt_1 ,$$

$$= \int_{a_0 \le t_2 \le t} \alpha_1(t_1, t_2)\tilde{A}(t_2) dt_2 = \alpha_2(t, a_0) \text{ , say.}$$

Similarly $\alpha_3(t,a_0) = \int_{a_0}^{t} \tilde{A}(t_1) \int_{a_0}^{t_1} \tilde{A}(t_2) \int_{a_0}^{t_2} \tilde{A}(t_3) dt_3 dt_2 dt_1$. Then (8) can be expressed more conveniently (by setting $\alpha_0(t,a_0) = \text{identity}$) as:

$$Y^*(t) = \tilde{Y}(t) = \sum_{n=0}^{\infty} \alpha_n(t,a_0)\tilde{Y}(a_0) . \tag{10}$$

As yet no special properties of Brownian motion, other than the definition of (6), were utilized. Let $Z(a_0) = C$ be a constant (nonstochastic) initial condition. (The existence and uniqueness of the solution of (3) rigorously holds if only $Z(a_0)$ is independent of $B(t)-B(a_0)$. However, this is not sufficient for the following analysis.) Also the continuity of $a(\cdot)$ and $b(\cdot)$ is not crucial. If $a(\cdot), b(\cdot)$ are integrable on T, then one sees easily that the bound in (9) holds (cf. also [24], Problem 1, p. 97), and the rest of the argument is valid. Thus one may state the following simple but important result.

Theorem 4.1. Let $T = [a_0, b_0)$ be a bounded interval and $\{B(t), t\in T\}$ be the Brownian motion. If $\{X(t), t\in T\}$ is a time series generated by (2) with $a(\cdot), b(\cdot)$ as the (Lebesgue) integrable real functions on T, then there is one and only one such series for each initial condition $X(a) = c_1$, $\dot{X}(a) = c_2$ where c_1, c_2 are real constants. The solution $Z(t) = \begin{bmatrix} \dot{X}(t) \\ X(t) \end{bmatrix}$ of (3) is given by (6) where Y is defined by (10). Moreover, $\{Z(t), t\in T\}$ is a vector Markov Gaussian time series, almost all of whose trajectories are continuous.

Proof. Because of the preceding discussion and calculations, only the proof of the last statement remains. Note that by the classical theory, the function $Y: T \to \mathbb{R}^4 - \{0\}$ is continuous and so also is Y^{-1}. By some well known properties of the simple stochastic integrals (cf. [13], Ch. 9) the integral in (6) is a continuous function of t with probability 1. It follows that the $Z(t)$-process has almost all continuous trajectories with values in $\mathbb{R}^2$.

By definition, the integral in (6) is the mean square limit of approximating sums of the form $\sum_{i=1}^{n} Y(t_i)(W(t_{i+1}) - W(t_i))$, where $a_0 = t_1 < t_2 < \cdots < t_n \leq t$. But since $Y(t_i)$ is nonstochastic and $W(t_{i+1}) - W(t_i)$ are independent Gaussian vectors, the above is Gaussian distributed. Hence its limit-in-mean (i.e. the integral) defines a Gaussian random vector. Since $Z(a_0) = C$, it follows from (6) that the $\{Z(t), t\in T\}$ is a vector Gaussian time series, having almost all continuous sample paths. Finally, (6) is true if $a < t$ is replaced by $t, t+h$ ($h>0$), a pair of points in T. Thus

$$Z(t) = \int_t^{t+h} Y(t+h)^{-1} Y(u) dW(u) + Y^{-1}(t+h) Y(t) Z(t) . \tag{11}$$

By the preceding definition of the integral, $Z(t)$ is determined by $Y(t)$ and $W(u)$ for $a_0 \le u \le t$ only. Since $W(\cdot)$ has independent increments, (11) is determined by $W(t_{i+1}) - W(t_i)$ for $t \le t_i < t_{i+1} \le t+h$, and is independent of $Z(u)$ for $a \le u < t$. This means for each interval $A \subset \mathbb{R}^2$, $t \ge t_0$,

$$P[Z_t \in A | Z_s, s \le t_0] = P[Z_t \in A | Z_0] , \tag{12}$$

with probability one. But this implies the Markovian property of the $Z(t)$, and completes the proof.

Remark. In general the Z-process need not be a martingale and the X-process need not be Markovian. Note that the cases $a(t) = a' + a''t$ and $b(t) = b' + b''t$ (the linear trend, and similarly the reciprocal trend if $T = [a_0, b_0], a_0 > 0$) is included in the above treatment. Another model treating seasonality (i.e. $a(t) = a_1 + a_2 \sin t$, $b(t) = b_1 + b_2 \cos t$) of the coefficients considered in ([5], p. 334) is also included. However, the latter case has some special properties worthy of a special treatment, but is not considered in the present paper. It is also clear that the above theorem is valid without real changes if the order of the equation is $n \ge 2$. In that case the $Y(t)$ and $A(t)$ are n-by-n matrices.

From (6) it is easy to calculate the mean and covariance of the $Z(t)$-vector series. In fact, if $m(t) = E(Z(t))$, $\Sigma(t,s) = \mathrm{Cov}(Z_s, Z_t)$, then from the condition that $Z(a_0) = C$, a constant, one has:

$$m(t) = Y^{-1}(t)Y(a_0)C \ . \tag{13}$$

and

$$\begin{aligned}\Sigma(s,t) &= E[(Z(s)-m(s))(Z(t)-m(t))^*] \\ &= E[(\int_{a_0}^{t} Y(t)^{-1}Y(u)dW(u))(\int_{a_0}^{s} Y(s)^{-1}Y(r)dW(r))^*] \\ &= \int_{a_0}^{s\wedge t} Y(t)^{-1}Y(u)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}Y^*(u)^{-1}Y^*(s)^{-1}du \ , \quad s\wedge t = \min(s,t) \ ,\end{aligned} \tag{14}$$

where an elementary property of the stochastic integral is used in the last line. If $D(t,u)$ is the first column of $Y(t)^{-1}Y(u)$, then (14) can be written as:

$$(\sigma_{ij}(s,t)) = \Sigma(s,t) = \int_{a_0}^{s\wedge t} D(t,u)D(s,u)^* du \ . \tag{15}$$

This is the 2-by-2 covariance matrix function of the vector $Z(t)$-series, which is (hermitian) positive definite for all $a_0 \leq s,t \leq b$. From this it follows immediately that the covariance function of the X-series ($r(s,t) = Cov(X_s, X_t) = \sigma_{22}(s,t)$) is given as:

$$r(s,t) = \int_{a_0}^{s\wedge t} (D(t,u)D(s,u)^*)_{22} du \tag{16}$$

where $(\ )_{22}$ is the second diagonal element of the matrix. If $d(s,u)$ is the second element of the vector $D(s,u)$, then (16) is just

$$r(s,t) = \int_{a_0}^{s\wedge t} d(s,u)\overline{d(t,u)} du \ . \tag{17}$$

Thus the time series determined by (2) is always of class (C).

It is interesting to note that the covariance function r of (17) is really the Green function of the homogeneous ordinary differential equation of (1) with zero initial conditions, i.e. $X(t) = \dot{X}(t) = 0$ at $t = a_0$. This useful result may be derived as follows.

Consider the homogeneous equation

$$L(X) = \frac{d^2X(t)}{dt^2} + a(t)\frac{dX(t)}{dt} + b(t)X(t) = 0 , \qquad (18)$$

where $X(a) = c_1$, $\dot{X}(a) = c_2$, as the initial data. If $U_i(a) = X^{(i-1)}(a)$ where $X^{(0)} = X$, $X^{(1)} = \dot{X} = \frac{dX}{dt}$, then by the standard theory (cf. [24]) the system $L(X) = 0$ and $U_1(a_0) = 0$ has a unique solution, say V_1. Similarly $L(X) = 0$ with $U_1(a_0) = 0$, $U_2(a_0) = 1$ has a unique solution V_2. If $G(\cdot,\cdot)$ is the Green function associated with $L(X) = 0$, $X(a_0) = 0 = \dot{X}(a_0)$, then the unique solution of (2) with the initial condition $Z(a) = C$ is given (cf. Ince [26], p. 257, and the discussion on the stochastic analog following (2)) by:

$$X(t) = \int_{a_0}^{t} G(t,u)dB(u) + C^*V(t) , \qquad (19)$$

where $V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$ and $C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$. Here G is continuous on $T \times T$ and differentiable (in general $(n-1)$ times for the n^{th} order equation), and $\frac{\partial G}{\partial t}$ is continuous in (t,u) for the range $a_0 \le t \le u \le b_0$. Moreover,

$$\frac{\partial G}{\partial t}(u+0,u) - \frac{\partial G}{\partial t}(u-0,u) = 1 .$$

Thus from (19) one has $m_1(t) = E(X(t)) = C^*V(t)$, and

$$r(s,t) = \mathrm{Cov}(X(s),X(t)) = \int_{a_0}^{s\wedge t} G(s,u)\overline{G(t,u)}du . \qquad (20)$$

A comparison of (19)-(20) with (16)-(17) shows that $G(s,u) = d(s,u)$ The form of this equation should also be compared with a result of Cramér ([27], p. 18) in terms of which $X(t)$ has multiplicity one. (For the definition and properties of the latter, cf., [27].)

The preceding work may be summarized in the following:

Theorem 4.2. The vector Gaussian Markov process $\{Z(t),t\in T\}$ of Theorem 4.1, has mean and covariance given by (13) and (15). The solution process $\{X(t),t\in T\}$ of (2) is given by (19) with its covariance function by (20) or (17), and it is of type (C) . If moreover, the Green kernel $G(\cdot,\cdot)$ of (20) satisfies an appropriate growth condition, in particular if $G(\cdot,u)$ is a.p. uniformly in u on compact subsets of T , then the solution process $\{X(t),t\in T\}$ is in class (C)$\cap$class (KF) .

Remark. If $a(t) = a_1 + a_2t$, $b(t) = b_1 + b_2t$, then the resulting solution process in the above theorem, while being in class (C) , will not be in class (KF) . In any case this solution has "multiplicity" one in the sense of ([27], p. 11). This problem and the work of Section II above show that the classes (C) and (KF) are not included in each other.

In view of the preceding result, it will be of considerable interest to analyze and extend the detailed treatment of [8] from the constant coefficient case to the general case which is a solution process of (2). Many of the results of [8] have non-trivial extensions to the present one. It needs and deserves a separate treatment and will not be considered at present. In the next paragraph a special problem is discussed for the properties of its correlogram, as it enables a better appreciation of the time series generated by (2).

(c) <u>A special case</u>. Let $a(t) = a$, $b(t) = b$, $t \in T$. Then $Y(\cdot)$ can be calculated explicitly in (10). For, $A(t) = A = \begin{bmatrix} a & b \\ -1 & 0 \end{bmatrix}$ so that $\alpha(t,t_0) = A^*(t-t_0)$, $\alpha_2(t,t_0) = A^{*2} \frac{(t-t_0)^2}{2}$ etc. and hence

$$Y^*(t) = \sum_{n=0}^{\infty} A^{*n} \frac{(t-a_0)^n}{n!} Y^*(a_0) = (\exp(t-a_0)A^*)Y^*(a_0) .$$

Consequently (6) can be simplified to

$$Z(t) = \int_{a_0}^{t} e^{(t-u)A} dW(u) + e^{(a_0-t)A} C , \tag{21}$$

where the facts that e^{uA} and e^{tA} commute and $Z(a_0) = C$ were used. From this, if $e^{(t-u)A} = \begin{pmatrix} d_1(t-u) & d_2(t-u) \\ d_3(t-u) & d_4(t-u) \end{pmatrix}$, then (21) yields

$$\dot{X}(t) = \int_{a_0}^{t} d_1(t-u)dB(u) + d_1(a_0-t)c_1 + d_2(a_0-t)c_2 ,$$

$$X(t) = \int_{a_0}^{t} d_3(t-u)dB(u) + d_3(a_0-t)c_1 + d_4(a_0-t)c_2 . \tag{22}$$

Comparing these equations with (19) one obtains $G(t,u) = G(t-u) = d_3(t-u)$, and $G'(t-u) = d_1(t-u)$ where G is the Green function associated with (2). (In the case of constant coefficients, G is only a function of the difference of its arguments.) The covariance is thus simplified to

$$r(s,t) = \int_0^{s \wedge t} G(s-u)\overline{G(t-u)}\,du \ . \tag{23}$$

Even now r is not a stationary covariance, and the solution process need not be in class (KF) without further conditions.

In the particular case that $a = 0$, $b>0$ in (21) with $T = [0,\infty)$, the covariance (23) can be simplified further. Then, $d_3(t) = \sin\sqrt{b}\,t$, and

$$r(s,t) = \int_0^{s \wedge t} \sin\sqrt{b}(t-u)\,\sin\sqrt{b}(s-u)\,du \ . \tag{24}$$

If $t = s+k$, then the correlation character $\tilde{\rho}_s(k) = r(s,s+k)/r(s,s)$ is reduced to:

$$\tilde{\rho}_s(k) = \frac{\int_0^s \sin\sqrt{b}(k+s-u)\,\sin\sqrt{b}(s-u)\,du}{\int_0^s \sin^2\sqrt{b}(s-u)\,du} \ , \qquad u \geq 0 \ ,$$

$$= \frac{2s\sqrt{b}\cos\sqrt{b}\,k + \sin\sqrt{b}\,k - \sin\sqrt{b}(k+2s)}{2s\sqrt{b} - \sin 2\sqrt{b}\,s} \ . \tag{25}$$

For large s, $\tilde{\rho}_s(k)$ therefore behaves as $\cos\sqrt{b}\,k$, and hence the approximate correlogram is essentially a harmonic curve. The above computation also shows that for <u>small</u> t, $r(t,t) = \int_0^t \sin^2\sqrt{b}\,u\,du \sim \frac{b}{3}t^3$. Similar estimates can be obtained for $\sigma_{ij}(t,t)$. Such calculations were carried out in ([8], p. 135) using different arguments, and both these results agree. The Gaussian character of $Z(t)$, when $a(t) = a$, $b(t) = b>0$, and various moments have long been established in [7], with classical techniques. It will thus be interesting to carry out the corresponding study for the case of variable $a(\cdot),b(\cdot)$, e.g., if these are periodic or more generally (real) analytic functions.

The result of Theorem 4.1 admits substantial extensions for vector valued processes and the coefficients can then be matrix valued integrable functions. Such an extension has been carried out by Goldstein [9] using suitable abstract methods. His results subsume all the previously known existence studies on the problem. The natural question here is to consider (and this appears more difficult) his results in classifying the solutions in the sense of Dym's work.

REFERENCES

1. Kampé de Fériet, J. and Frenkiel, F. N., "Correlations and Spectra for Nonstationary Random Functions," Math. Computation 16, 1-21, 1962.

2. Parzen, E., "Spectral Analysis of Asymptotically Stationary Time Series," Bull. Inst. Int'l. Statist. 39 (livraison 2), 87-103, 1962.

3. Bhagavan, C. S. K., Nonstationary Processes, Spectra and some Ergodic Theorems, Andhra Univ. Press, Waltair, India, 1974.

4. Cramér, H., "A Contribution to the Theory of Stochastic Processes," Proc. 2d. Berkeley Symp. Math. Statist. and Prob., 329-339, 1951.

5. Hurwicz, L., "Variable Parameters in Stochastic Processes: Trend and Seasonality," in Statistical Inference in Dynamic Economic Models, 329-344, Wiley, 1950.

6. Hartley, H. O., "Second Order Autoregressive Schemes with Time-Trending Coefficients," J. Roy. Statist. Soc. (B) 14, 229-233, 1952.

7. Chandrasekhar, S., "Stochastic Problems in Physics and Astronomy," Rev. Mod. Phys. 15, 1-89, 1943.

8. Dym, H., "Stationary Measures for the Flow of a Linear Differential Equation Driven by White Noise," Trans. Amer. Math. Soc. 123, 130-164, 1966.

9. Goldstein, J. A., "An Existence Theorem for Linear Stochastic Differential Equations," J. Diff. Eqs. 3, 78-87, 1967.

10. Rao, M. M., "Representation Theory of Multidimensional Generalized Random Fields," Multivariate Analysis II, 411-436, Academic Press, 1969.

11. Yoshizawa, T., Stability Theory and the Existence of Periodic Solutions and Almost Periodic Solutions, Applied Math. Series No. 14, Springer, 1975.

12. Besicovitch, A. S., Almost Periodic Functions, Cambridge Univ. Press, 1932.

13. Doob, J. L., Stochastic Processes, Wiley, 1953.

14. Cramér, H., Random Variables and Probability Distributions, Cambridge Univ. Press, 1937.

15. Rao, M. M., Specification of some Stochastic Models, unpublished M. Sc. Thesis, Madras Univ., 1955.

16. Jordan, C., Calculus of Finite Differences, (2d. ed.) Chelsea, 1950.

17. Mann, H. B. and Wald, A., "On the Statistical Treatment of Linear Stochastic Difference Equations," Econometrica 11, 173-220, 1943.

18. Rao, M. M., "Inference in Stochastic Processes - II," Z. Wahrs. 4, 317-335, 1966.

19. Rao, M. M., "Consistency and Limit Distributions of Estimators of Parameters in Explosive Stochastic Difference Equations," Ann. Math. Statist. 32, 195-218, 1961.

20. Anderson, T. W., The Statistical Analysis of Time Series, Wiley, 1971.

21. Marsaglia, G., "Iterated Limits and the Central Limit Theorem for Dependent Random Variables," Proc. Amer. Math. Soc. 5, 987-991, 1954.

22. Parzen, E., "A Central Limit Theorem for Multilinear Stochastic Processes," Ann. Math. Statist. 28, 252-256, 1957.

23. Nagabhushanam, K., "The Primary Process of a Smoothing Relation," Ark. für Mat. 1, 421-488, 1951.

24. Coddington, E. A. and Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill, 1955.

25. Doob, J. L., "The Brownian Movement and Stochastic Equations," Ann. Math. 43, 351-369, 1942.

26. Ince, E. L., Ordinary Differential Equations, London, 1927.

27. Cramér, H., Structural and Statistical Problems for a Class of Stochastic Processes, Princeton Univ. Press, 1971.

☆ U. S. GOVERNMENT PRINTING OFFICE: 1977 — 757-001/260